

DESEMPENHO DE INTERVALOS DE CONFIANÇA PARA A MÉDIA DE POPULAÇÕES POISSON AVALIADO POR SIMULAÇÃO MONTE CARLO

Cássio Bueno Vilela CORRÊA¹
Daniel Furtado FERREIRA²

- RESUMO: Várias aproximações do intervalo de confiança para o parâmetro λ de uma população Poisson existem na literatura. O presente trabalho foi realizado com o objetivo principal de avaliar, por simulação, as propriedades de alguns procedimentos assintóticos para o intervalo de confiança do parâmetro λ da Poisson. Foram estudados três intervalos assintóticos e o intervalo exato. As simulações foram implementadas no *software* R. Como objetivos específicos almejou-se obter uma solução explícita para o intervalo de Hall (1982) e generalizar para o caso de $n > 1$, o intervalo de Bégaud et al. (2005). Uma solução explícita para o intervalo de Hall (1982) foi apresentada como proposto. O intervalo de confiança aproximado de Bégaud et al. (2005) foi generalizado para amostras com $n \geq 1$. Os três intervalos assintóticos apresentam desempenho equivalente ao intervalo exato em relação a probabilidade de cobertura, a partir de $\lambda \geq 5$ e $n > 1$. No caso de $n = 1$ e $\lambda \geq 5$, o intervalo de Bégaud et al. (2005) teve melhor desempenho. Em relação ao comprimento dos intervalos, para $\lambda \geq 5$ e $n > 1$, os desempenhos dos intervalos de confiança seguiram a seguinte ordem; do melhor (menor comprimento) para o pior (maior comprimento): Intervalo de Hall (1982), Ferreira (2009), Bégaud et al. (2005) e exato. Como em pequenas amostras, para pequenos valores de λ , as aproximações tiveram baixo desempenho e o valor de λ é desconhecido, recomenda-se a utilização do intervalo exato, exceto na derivação de algum resultado em que a teoria exata seja impossível de ser aplicada.
- PALAVRAS-CHAVE: Aproximação normal da Poisson; comprimento de intervalo; probabilidade de cobertura; Monte Carlo.

¹Universidade Federal de Lavras – UFLA, Programa de Pós-graduação em Estatística e Experimentação Agronômica, CEP: 37200-000, Lavras, MG, Brasil. E-mail: ccorrea@computacao.ufla.br

²Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas – DEX, CEP: 37200-000, Lavras, MG, Brasil. E-mail: danielff@dex.ufla.br

1 Introdução

Ao se estudar uma população, não é possível medir-se todas as observações, em razão do custo elevado e do tempo requerido. Assim, se torna necessário realizar procedimentos de amostragem para se estimar algum parâmetro. Os processos de amostragem e de estimação dependem das características de cada população. As populações possuem características peculiares que as determinam, como tipo de variável, se discreta ou contínua, tipo de disposição física, dependência espacial entre as observações, entre outras. As populações de interesse desse trabalho são aquelas caracterizadas por um fenômeno que ocorre de forma aleatória e cujos dados são provenientes de contagens dos eventos de interesse que ocorrem por unidade de área, volume ou tempo (FERREIRA, 2009, p.113), se os eventos ocorrem em intervalos de tempo, espaço ou volume fixos, de forma aleatória, com taxa constante e independentes da ocorrência do último evento. Essas populações são denominadas de populações Poisson. São exemplos de variáveis aleatórias Poisson a contagem do número de formigueiros numa área, o número de bactérias por unidade de área em uma lâmina, número de telefonemas em uma central, entre outros, desde que obedecem as condições anteriormente colocadas. O parâmetro dessa população refere-se a média da ocorrência desse fenômeno por unidade de área, volume e tempo, sendo denotado por λ . Assim, uma variável aleatória X com distribuição Poisson tem a seguinte função de probabilidade

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad (1)$$

para $x = 0, 1, 2, \dots$

Ao se estimar o parâmetro λ da Poisson é comum utilizar-se de intervalos de confiança e fixar o coeficiente de confiança, de que o intervalo contenha o verdadeiro valor do parâmetro desconhecido, em $1 - \alpha = 90, 95$ e 99% . Manter o valor de confiança fixado em elevados valores tem custos, que devem ser minimizados com a proposição de novos métodos que sejam mais eficientes. O custo, em geral, é reflexo dos maiores tamanhos amostrais requeridos para que o intervalo possua elevado coeficiente de confiança e mantenha-se útil, ou seja, com amplitude pequena.

Existem muitas alternativas para se obter intervalos de confiança para o parâmetro λ da distribuição Poisson. Entre elas, está o intervalo exato. Portanto, na distribuição Poisson, o intervalo de confiança exato (GARWOOD, 1936) para λ (μ), com coeficiente de confiança de pelo menos a $1 - \alpha$ é dado por

$$IC_{1-\alpha}(\lambda = \mu) : \left[\lambda_I = \frac{1}{2n} \chi_{\alpha/2; \nu=2k}^2; \lambda_S = \frac{1}{2n} \chi_{1-\alpha/2; \nu=2k+2}^2 \right], \quad (2)$$

em que $\chi_{\alpha/2; \nu=2k}^2$ e $\chi_{1-\alpha/2; \nu=2k+2}^2$ são os quantis $\alpha/2$ e $1 - \alpha/2$, da distribuição qui-quadrado com $\nu = 2k$ e $2k+2$ graus de liberdade, respectivamente; $k = \sum_{i=1}^n X_i$, sendo X_1, X_2, \dots, X_n uma amostra aleatória de tamanho n . Se $k = 0$, então $\lambda_I = 0$.

É possível obter aproximações normais assintóticas para o intervalo de confiança de λ , sendo interessante recordar que só se tem boa aproximação normal, se λ for grande (maior que 7). Pode-se derivar um intervalo normal aproximado considerando que $V(\bar{X}) = nV(X_i/n) = \lambda/n$. Portanto, usando-se o estimador $\hat{\lambda}/n$ dessa variância, a aproximação é dada, segundo Ferreira (2009) por:

$$IC_{1-\alpha}(\mu) : \left[\hat{\lambda} + Z_{\alpha/2} \sqrt{\frac{\hat{\lambda}}{n}}; \hat{\lambda} + Z_{1-\alpha/2} \sqrt{\frac{\hat{\lambda}}{n}} \right] \\ = \left[\bar{X} + Z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}}; \bar{X} + Z_{1-\alpha/2} \sqrt{\frac{\bar{X}}{n}} \right],$$

em que $Z_{\alpha/2}$ e $Z_{1-\alpha/2}$ são os quantis $\alpha/2$ e $1 - \alpha/2$, respectivamente, da normal padrão e $\hat{\lambda} = \bar{X} = \sum_{i=1}^n X_i/n$.

Esta aproximação é ruim se λ é pequeno (menor que 7) (FERREIRA, 2009) para n pequeno, sendo deste modo interessante avaliar outros métodos que forneçam intervalos de confiança que corresponda a uma melhor aproximação, quando comparado com o intervalo de confiança exato dado pela expressão (2).

Outra aproximação é devido a Blyth e Hutchinson (1961), em que os autores consideram uma única unidade amostral, ou seja, considerando $n = 1$. Nesse caso, $\bar{X} = X$, que é a única observação amostral considerada. Portanto, o intervalo de $100(1 - \alpha)\%$ de confiança é

$$IC_{1-\alpha}(\mu) : \frac{1}{2} \left[Z_{1-\alpha/2}^2 + 2X \pm Z_{1-\alpha/2} \left(Z_{1-\alpha/2}^2 + 4X \right)^{1/2} \right]. \quad (3)$$

Ferreira (2009) derivou essa expressão para o caso de n observações amostrais, a qual é dada por

$$IC_{1-\alpha}(\mu) : \frac{1}{2} \left[\frac{Z_{1-\alpha/2}^2}{n} + 2\bar{X} \pm \frac{Z_{1-\alpha/2}}{\sqrt{n}} \left(\frac{Z_{1-\alpha/2}^2}{n} + 4\bar{X} \right)^{1/2} \right]. \quad (4)$$

Patil e Kulkarni (2012) avaliaram dezenove métodos para obtenção de intervalos de confiança para o parâmetro λ da Poisson. Em todos os casos foi considerada a situação de $n = 1$. Um grupo de intervalos foi considerado uniformemente ótimo em relação ao domínio do parâmetro λ . Dentre os métodos desse grupo, o intervalo de Bégaud et al. (2005) foi escolhido nesse trabalho pelo critério da parcimônia. Esse intervalo é

$$IC_{1-\alpha}(\mu) : \left[\left(\sqrt{X + 0,02} + \frac{Z_{\alpha/2}}{2} \right)^2 ; \left(\sqrt{X + 0,96} + \frac{Z_{1-\alpha/2}}{2} \right)^2 \right]. \quad (5)$$

para $n = 1$, sendo X a variável aleatória Poisson.

Uma outra aproximação foi apresentada por Hall (1982) e não incluída no trabalho de comparação de intervalos de confiança para λ apresentado por Patil e Kulkarni (2012). Esse método (HALL, 1982) não apresenta uma solução explícita para se obter os limites inferior e superior de confiança do intervalo, sendo necessário utilizar algum método numérico. Em decorrência de o trabalho de Patil e Kulkarni (2012) não apresentar estudos para valores de n superiores a 1 e não considerar algumas das aproximações mencionadas, surgiu a ideia de desenvolver esse trabalho.

O presente trabalho foi realizado com o objetivo principal de avaliar, por simulação, as propriedades de alguns procedimentos assintóticos para o intervalo de confiança do parâmetro λ da Poisson. Foram estudados três intervalos assintóticos (HALL, 1982; BÉGAUD et al., 2005; FERREIRA, 2009) e o intervalo exato. As simulações foram implementadas no *software* R (R Core Team, 2013). Como objetivos específicos almejou-se obter uma solução explícita para o intervalo de Hall (1982) e generalizar para o caso de $n > 1$ o intervalo de Bégaud et al. (2005).

2 Metodologia

A presente proposta teve intenção de avaliar por simulação Monte Carlo o desempenho de três opções para o intervalo de confiança para o parâmetro $\lambda = \mu$ da distribuição Poisson. O desempenho foi medido por intermédio das probabilidades de cobertura e dos comprimentos dos intervalos considerados.

2.1 Intervalos de confiança estudados

Considere uma amostra aleatória X_1, X_2, \dots, X_n de observações independentes e identicamente distribuídas como uma Poisson, cuja função de probabilidade é dada em (1). Foi considerado o estimador pontual de máxima verossimilhança de λ , dado por

$$\hat{\lambda} = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}.$$

O primeiro intervalo de confiança adotado foi o exato, cuja expressão foi apresentada em (2). Esse intervalo foi utilizado para fins de comparação, servindo de referência para as demais aproximações consideradas.

Foram consideradas três aproximações normais. A primeira foi apresentada na expressão (4), sendo desenvolvida por Ferreira (2009). A segunda aproximação normal foi apresentada em Hall (1982). A solução para esse intervalo de confiança não foi explicitada pelo autor. Portanto, a solução da inequação

$$P\left(\sqrt{n}(\hat{\lambda} - \lambda) \leq \sqrt{\hat{\lambda}}Z_p + \frac{(1 - Z_p^2)/3}{\sqrt{n}}\right) \leq \Phi(Z_p) + O(n^{-1}), \quad (6)$$

em relação a λ será obtida para se construir um intervalo de $100(1 - \alpha)\%$ de confiança. Deve-se chamar a atenção que Z_p refere-se ao quantil $100p\%$ da distribuição normal padrão, ou seja, $Z_p = \Phi^{-1}(p)$, em que $\Phi^{-1}(p)$ é a função de distribuição inversa da normal padrão. Assim, para um intervalo de $100(1 - \alpha)\%$ dois valores de p devem ser considerados: $p = \alpha/2$ e $1 - \alpha/2$, para os limites inferior e superior do intervalo de confiança, respectivamente. A solução dessa expressão foi considerada como resultado do presente trabalho. A terceira aproximação é a de Bégau et al. (2005) dada pela expressão (5). A expressão (5) foi também estendida para os casos em que $n > 1$.

2.2 Avaliação do desempenho dos intervalos

Para a avaliação do desempenho dos intervalos de confiança considerados, foram utilizados o comprimento e a probabilidade de cobertura. Sendo LI e LS os limites inferior e superior, respectivamente, de um dado intervalo, então o comprimento é definido por

$$L = LS - LI. \quad (7)$$

Como foram realizadas um número grande de simulações, N , então para cada intervalo em cada simulação tem-se um comprimento. Para fins de resumo e comparação, foram considerados os comprimentos médios das N simulações para os quatro intervalos estudados.

A probabilidade de cobertura refere-se a probabilidade de que um intervalo contenha o verdadeiro valor do parâmetro. Sendo $I_i(IC_k)$ uma função indicadora de que o k -ésimo intervalo de confiança, $k = 1, 2, 3, 4$, contenha o verdadeiro valor paramétrico λ na i -ésima simulação, $i = 1, 2, \dots, N$, de uma dada configuração, então a probabilidade de cobertura do k -ésimo intervalo foi estimada por

$$\widehat{PC}_k = \frac{\sum_{i=1}^N I_i(IC_k)}{N}. \quad (8)$$

Para a realização das simulações foram geradas N amostras de tamanho n de uma distribuição Poisson com parâmetro λ , cuja função de probabilidade foi apresentada na expressão (1). Assim, em cada uma das N simulações uma amostra aleatória dada por X_1, X_2, \dots, X_n foi gerada da distribuição Poisson. Para isso foi utilizado a função `rpois(n, lambda)` do programa R (R Core Team, 2013). O algoritmo usado na geração de realizações de variáveis aleatórias Poisson da função `rpois` foi proposto por Ahrens e Dieter (1982), baseado no método de aceitação/rejeição. Gerada a amostra na i -ésima simulação Monte Carlo, os quatro intervalos apresentados na subseção 2.1 foram computados. O processo foi repetido para $i = 1, 2, \dots, N$. Para cada intervalo, o comprimento médio e a probabilidade de cobertura, foram computados ao final das N simulações.

Os resultados foram tabulados ou apresentados na forma gráfica e o intervalo de menor comprimento e probabilidade de cobertura próximo ao coeficiente

de confiança nominal de $100(1 - \alpha)\%$ foi considerado ótimo. Para que esses resultados tenham uma representatividade ampla, vários cenários (configurações) foram considerados, em relação ao tamanho da amostra n , valor do parâmetro λ e coeficiente de confiança $100(1 - \alpha)\%$. Além disso, foram consideradas $N = 10.000$ simulações Monte Carlo. Os valores de n considerados nas simulações foram: $n = 1, 5, 10(5)30, 50, 100, 1.000$. Os valores de λ foram: $\lambda = 0,01, 0,10, 0,50, 1, 5$ e 7 . A escolha desses valores ocorreu em virtude de se ter situações de pequenas e grandes amostras e de valores paramétricos para os quais as aproximações assintóticas, em princípio, tenham desempenhos ruins (pequenos valores) e bons (grandes valores). Os valores do coeficiente de confiança utilizados nesse trabalho foram 95% e 99% .

3 Resultados e discussão

Para explicitar a expressão de Hall (1982), foram considerados, em (6), os valores de p iguais a $\alpha/2$ e $1 - \alpha/2$. Assim, para $p = \alpha/2$ o valor de $\Phi(Z_p)$ é $\Phi(Z_{\alpha/2}) = \alpha/2$. Portanto,

$$P\left(\sqrt{n}(\hat{\lambda} - \lambda) \leq \sqrt{\hat{\lambda}}Z_{\alpha/2} + \frac{(1 - Z_{\alpha/2}^2)/3}{\sqrt{n}}\right) \leq \alpha/2$$

$$P\left(\lambda \geq \hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}}Z_{\alpha/2} - \frac{1 - Z_{\alpha/2}^2}{3n}\right) \leq \alpha/2.$$

Se for observada a expressão anterior em detalhe, verifica-se que

$$\hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}}Z_{\alpha/2} - \frac{1 - Z_{\alpha/2}^2}{3n}$$

corresponde ao limite superior do intervalo de confiança almejado, uma vez que se trata do quantil superior $100\alpha/2\%$. Considerando $p = 1 - \alpha/2$, pode-se obter o limite inferior do intervalo da mesma forma. Assim, substituindo-se p por $1 - \alpha/2$ na expressão (6), após algumas simplificações, obtém-se:

$$P\left(\lambda \geq \hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}}Z_{1-\alpha/2} - \frac{1 - Z_{1-\alpha/2}^2}{3n}\right) \leq 1 - \alpha/2,$$

em que

$$\hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}}Z_{1-\alpha/2} - \frac{1 - Z_{1-\alpha/2}^2}{3n}$$

corresponde ao limite inferior do intervalo de $100(1 - \alpha)\%$ de confiança, por se tratar do quantil superior $100(1 - \alpha/2)\%$. Logo,

$$IC_{1-\alpha}(\mu) : \left[\hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}} Z_{1-\alpha/2} - \frac{1 - Z_{1-\alpha/2}^2}{3n}; \hat{\lambda} - \sqrt{\frac{\hat{\lambda}}{n}} Z_{\alpha/2} - \frac{1 - Z_{\alpha/2}^2}{3n} \right]. \quad (9)$$

A generalização do intervalo (5) para qualquer n , tal que $n \geq 1$, foi realizada. Inicialmente, considerou-se que

$$X = \sum_{i=1}^n X_i.$$

A partir dessa definição, o intervalo de confiança (5) irá representar o intervalo para o parâmetro $n\lambda$, pois a soma de variáveis Poisson independentes e identicamente distribuídas é uma Poisson, ou seja, $\sum_{i=1}^n X_i \sim \text{Poi}(n\lambda)$. Portanto, dividindo-se os limites do intervalo (5) por n e substituindo X por $\sum_{i=1}^n X_i$, o intervalo resultante corresponderá ao intervalo de confiança para o parâmetro λ . Logo, para o limite inferior(LI), tem-se:

$$\begin{aligned} LI &= \frac{\left(\sqrt{\sum_{i=1}^n X_i + 0,02 + \frac{Z_{\alpha/2}}{2}} \right)^2}{n} \\ &= n \left(\frac{\sqrt{\sum_{i=1}^n X_i + 0,02}}{n} + \frac{Z_{\alpha/2}}{2n} \right)^2 \\ &= n \left(\frac{\sqrt{\bar{X} + 0,02/n}}{n} + \frac{Z_{\alpha/2}}{2n} \right)^2. \end{aligned}$$

Adotando-se um procedimento similar para o limite superior, obtém-se o seguinte intervalo:

$$IC_{1-\alpha}(\mu) : \left[n \left(\frac{\sqrt{\bar{X} + 0,02/n}}{n} + \frac{Z_{\alpha/2}}{2n} \right)^2 ; n \left(\frac{\sqrt{\bar{X} + 0,96/n}}{n} + \frac{Z_{1-\alpha/2}}{2n} \right)^2 \right].$$

O desempenho dos quatro intervalos de confiança foram avaliados por Monte Carlo em 10.000 repetições em função do tamanho da amostra e do parâmetro λ .

Na Tabela 1 são apresentadas as probabilidades de coberturas estimadas nas 10.000 simulações Monte Carlo em função do n , para $\lambda = 0,01, 0,10, 0,50, 1,00, 5,00$ e $7,00$, considerando $\alpha = 0,01$. Observa-se que quanto menor for o valor de λ , pior é o desempenho das três aproximações, principalmente para menores tamanhos de amostra. Em alguns casos, as probabilidades de cobertura atingem valores extremamente pequenos, quando comparados com o valor nominal de 99%. O intervalo exato, como era esperado, apresentou probabilidades de cobertura iguais ou superiores ao valor nominal de 99% em todos os casos considerados, desprezando-se o erro de Monte Carlo associado.

Table 1 - Probabilidade de cobertura dos quatro intervalos estudados (IC1 - Exato; IC2 - Ferreira (2009); IC3 - Hall (1982); IC4 - Bégau (2009)) em $N = 10.000$ para diferentes valores de n , considerando $\lambda = 0,01, 0,10, 0,50, 1,00, 5,00$ e $7,00$ com $\alpha = 0,01$

n	$\lambda = 0,01$				$\lambda = 0,10$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	1,0000	0,9881	0,9881	0,0000	0,9960	0,9084	0,9084	0,0916
5	0,9995	0,9494	0,9494	0,0005	0,9986	0,9869	0,9986	0,3933
10	0,9960	0,9078	0,9078	0,0919	0,9958	0,9807	0,9958	0,6308
15	0,9994	0,9905	0,8698	0,1296	0,9959	0,9817	0,9959	0,9992
20	0,9987	0,9824	0,8235	0,1752	0,9958	0,9843	0,8639	0,9985
25	0,9978	0,9753	0,8012	0,2213	0,9974	0,9886	0,9125	0,9991
30	0,9970	0,9648	0,7750	0,2570	0,9966	0,9879	0,9498	0,9988
50	0,9988	0,9856	0,9988	0,3903	0,9984	0,9873	0,9902	0,9984
100	0,9967	0,9802	0,9967	0,6334	0,9944	0,9908	0,9805	0,9944
1000	0,9949	0,9925	0,9889	0,9949	0,9935	0,9928	0,9866	0,9935
n	$\lambda = 0,50$				$\lambda = 1,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	0,9984	0,9864	0,9984	0,3913	0,9963	0,9805	0,9963	0,6295
5	0,9958	0,9852	0,9124	0,9992	0,9977	0,9862	0,9900	0,9977
10	0,9979	0,9865	0,9912	0,9979	0,9948	0,9945	0,9870	0,9948
15	0,9913	0,9899	0,9913	0,9933	0,9948	0,9872	0,9901	0,9948
20	0,9941	0,9924	0,9861	0,9941	0,9904	0,9900	0,9904	0,9931
25	0,9932	0,9878	0,9895	0,9930	0,9933	0,9871	0,9900	0,9933
30	0,9936	0,9857	0,9876	0,9936	0,9922	0,9920	0,9891	0,9922
50	0,9933	0,9880	0,9900	0,9933	0,9925	0,9913	0,9896	0,9925
100	0,9901	0,9898	0,9882	0,9901	0,9923	0,9902	0,9906	0,9923
1000	0,9926	0,9918	0,9920	0,9920	0,9894	0,9890	0,9890	0,9898
n	$\lambda = 5,00$				$\lambda = 7,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	0,9981	0,9866	0,9909	0,9981	0,9968	0,9869	0,9905	0,9968
5	0,9934	0,9887	0,9897	0,9934	0,9909	0,9909	0,9872	0,9909
10	0,9912	0,9901	0,9888	0,9912	0,9927	0,9916	0,9912	0,9931
15	0,9898	0,9896	0,9880	0,9898	0,9894	0,9894	0,9870	0,9915
20	0,9926	0,9901	0,9910	0,9926	0,9921	0,9909	0,9911	0,9921
25	0,9918	0,9894	0,9903	0,9918	0,9922	0,9919	0,9914	0,9922
30	0,9916	0,9906	0,9896	0,9902	0,9898	0,9896	0,9889	0,9898
50	0,9929	0,9911	0,9917	0,9929	0,9903	0,9904	0,9896	0,9903
100	0,9906	0,9886	0,9893	0,9893	0,9899	0,9900	0,9891	0,9899
1000	0,9905	0,9906	0,9904	0,9905	0,9899	0,9898	0,9897	0,9898

Para $\lambda = 0,01$, que pode ser considerado o pior cenário simulado para as aproximações, alguns dos métodos apresentaram cobertura muito pequenas e um deles registrou valor zero ou muito próximo de zero com $n \leq 5$, que foi o de Bégau. Este mesmo intervalo possui probabilidades de cobertura inferior ao valor nominal, quando $n\lambda$ é inferior a um valor próximo de 1. Assim, nesse caso, mesmo com $n > 100$ (dados não apresentados), a probabilidade de cobertura foi inferior ao valor nominal de 99%.

Os intervalos de Hall e Ferreira apresentaram quedas de coberturas com o aumento dos tamanhos amostrais, para λ pequenos. Após atingir um certo limite, a cobertura volta a crescer e atinge os valores nominais e novamente há uma queda com o aumento do n . Esse processo se repete em vários ciclos (dados completos não apresentados). O intervalo de Hall apresenta quedas mais abruptas, porém as fases de recuperação da cobertura são mais longas. Por exemplo, a cobertura do intervalo de Hall é de 99%, desprezando-se o erro de Monte Carlo, com $n = 1$ para $\lambda = 0,01$. Essa cobertura é inferior a 99% para n variando de 2 a 30, sendo que com $n = 30$ o seu valor é mínimo. Para $n = 30$ até 186 o valor de cobertura é no mínimo igual ao valor nominal de 99%, vindo a cair novamente com $n = 187$. A partir desse ponto, as coberturas voltam a subir com o aumento do n , mas permanecem por um certo tempo abaixo do valor nominal. Já o intervalo de Ferreira apresenta quedas com o aumento do n , partindo do valor nominal. Essas quedas diminuem de intensidade quando os tamanhos amostrais aumentam. Para valores de $\lambda \geq 0,10$, as taxas de cobertura do intervalo de confiança de Ferreira são praticamente iguais aos valores nominais. Para $\lambda = 1$, somente os intervalos de Bégau com $n = 1$ e o de Hall com n entre 2 e 4 (dados não apresentados) possuem cobertura inferior a 99%.

As probabilidades de cobertura de todos os 4 intervalos são iguais ou superiores ao valor nominal de 99% para todos os n , nas situações que caracterizam os casos em que os valores de λ são considerados grandes ($\lambda = 5$ e 7). Isso aconteceu, em geral, também para todos os casos em que $\lambda > 1$ (dados não apresentados).

Na Tabela 2 estão apresentadas as probabilidades de cobertura para os 4 intervalos considerados nesse trabalho, para $\alpha = 0,05$. Essas 4 situações são as mesmas da Tabela 1, exceto pelo valor de α . Os intervalos tiveram um padrão de resposta semelhante ao caso anteriormente mencionado. A recuperação do intervalo de Bégau em relação a probabilidade de cobertura se deu em torno de $n = 70$ para $\lambda = 0,01$, e para $n = 8$ com $\lambda = 0,10$ e para $n = 2$ com $\lambda = 0,50$. O intervalo de Hall apresenta resultados inesperados com determinadas combinações de n e λ , para $\lambda \leq 1$. O intervalo de Ferreira manteve seu padrão de resposta que havia sido observado para $\alpha = 0,01$. O intervalo exato apresentou probabilidades de cobertura sempre iguais ou superiores ao valor nominal, como esperado.

As probabilidades de cobertura dos 4 intervalos para $\lambda = 5$ e 7 , com $\alpha = 0,05$, permitiram verificar que os intervalos apresentaram desempenhos equivalentes para todos os valores de n e foram considerados exatos. Esse tipo de comportamento já é esperado pela teoria, uma vez que quando o valor de λ cresce, as aproximações normais assintóticas são melhores. Em estudos da distribuição Poisson, muitos autores, como, por exemplo, (FERREIRA, 2009), apontam que quando, na prática,

Table 2 - Probabilidade de cobertura dos quatro intervalos estudados (IC1 - Exato; IC2 - Ferreira (2009); IC3 - Hall (1982); IC4 - Bégau (2009)) em $N = 10.000$ para diferentes valores de n , considerando $\lambda = 0,01, 0,10, 0,50, 1,00, 5,00$ e $7,00$ com $\alpha = 0,05$

n	$\lambda = 0,01$				$\lambda = 0,10$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	0,9895	0,9895	0,9999	0,0090	0,9956	0,9098	0,9956	0,0919
5	0,9990	0,9483	0,9990	0,0485	0,9850	0,9127	0,9850	0,3824
10	0,9952	0,9034	0,9952	0,0896	0,9826	0,9189	0,6228	0,9827
15	0,9997	0,9863	0,9897	0,1250	0,9805	0,9329	0,7513	0,9812
20	0,9942	0,9842	0,9995	0,1835	0,9841	0,9477	0,8473	0,9829
25	0,9983	0,9755	0,9983	0,2216	0,9837	0,9569	0,8997	0,9867
30	0,9664	0,9615	0,9965	0,2544	0,9869	0,9659	0,9386	0,9883
50	0,9864	0,9091	0,9864	0,3853	0,9801	0,9632	0,9460	0,9788
100	0,9808	0,9259	0,6100	0,9816	0,9779	0,9650	0,9601	0,9743
1000	0,9749	0,9626	0,9543	0,9743	0,9536	0,9485	0,9471	0,9558
n	$\lambda = 0,50$				$\lambda = 1,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	0,9867	0,9086	0,9867	0,3831	0,9800	0,9166	0,6157	0,9809
5	0,9846	0,9586	0,9018	0,9863	0,9788	0,9601	0,9462	0,9786
10	0,9809	0,9643	0,9468	0,9779	0,9769	0,9626	0,9584	0,9768
15	0,9590	0,9361	0,9196	0,9612	0,9637	0,9511	0,9464	0,9598
20	0,9755	0,9623	0,9560	0,9742	0,9567	0,9440	0,9396	0,9595
25	0,9671	0,9548	0,9473	0,9698	0,9503	0,9391	0,9326	0,9550
30	0,9594	0,9475	0,9399	0,9625	0,9585	0,9496	0,9469	0,9532
50	0,9514	0,9404	0,9350	0,9534	0,9539	0,9494	0,9448	0,9545
100	0,9512	0,9421	0,9420	0,9559	0,9501	0,9443	0,9444	0,9542
1000	0,9517	0,9491	0,9492	0,9502	0,9531	0,9487	0,9505	0,9516
n	$\lambda = 5,00$				$\lambda = 7,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	0,9804	0,9631	0,9478	0,9800	0,9791	0,9642	0,9569	0,9812
5	0,9554	0,9451	0,9400	0,9533	0,9560	0,9472	0,9456	0,9579
10	0,9560	0,9495	0,9451	0,9505	0,9632	0,9563	0,9561	0,9583
15	0,9518	0,9459	0,9518	0,9539	0,9588	0,9557	0,9542	0,9621
20	0,9523	0,9470	0,9455	0,9526	0,9540	0,9493	0,9485	0,9571
25	0,9515	0,9456	0,9449	0,9459	0,9482	0,9438	0,9438	0,9503
30	0,9538	0,9538	0,9500	0,9534	0,9541	0,9503	0,9500	0,9545
50	0,9519	0,9456	0,9482	0,9509	0,9513	0,9489	0,9476	0,9498
100	0,9514	0,9485	0,9484	0,9507	0,9499	0,9476	0,9471	0,9485
1000	0,9484	0,9477	0,9470	0,9494	0,9489	0,9465	0,9477	0,9522

λ é superior ou igual a 5, as aproximações normais são adequadas para modelar a distribuição Poisson. Os dois casos apresentados possuem $\lambda \geq 5$ e, portanto, atendem a esse requisito, o que pode ter sido a causa do excelente desempenho de todas as aproximações consideradas.

Na Figura 1 estão apresentadas as probabilidades de cobertura para os 4 intervalos estudados, considere $\alpha = 0,01$, em função de λ , considerando diferentes valores de n . O intervalo de Bégau apresentou cobertura inferior ao valor nominal de 99% para pequenos valores de λ . Com o aumento dos valores de λ , a probabilidade de cobertura se aproximou do valor nominal, como o esperado e

comentado anteriormente. Considerando o caso de $n = 1$ e $\lambda = 0,01$, a probabilidade de cobertura foi inferior ao valor nominal de 99%, sendo que a recuperação, em relação a probabilidade de cobertura, se deu em torno de $\lambda = 1,75$. Ao se aumentar os tamanhos amostrais, houve uma recuperação da probabilidade de cobertura com um valor inferior de λ . Assim, considerando-se $n = 10$, a recuperação do intervalo de Bégaud se deu para $\lambda = 0,66$. O intervalo de Hall apresentou quedas abruptas na probabilidade de cobertura, principalmente para pequenos valores de n ou pequenos valores de λ . Para esse intervalo, com $n = 1$, a probabilidade de cobertura foi inferior ao valor nominal de 99%, principalmente, para valores de λ entre 2 e 4,7. Houve uma recuperação após esse limite de 4,7. Com tamanhos amostrais maiores, todos os intervalos estudados apresentaram cobertura de pelo menos 99%.

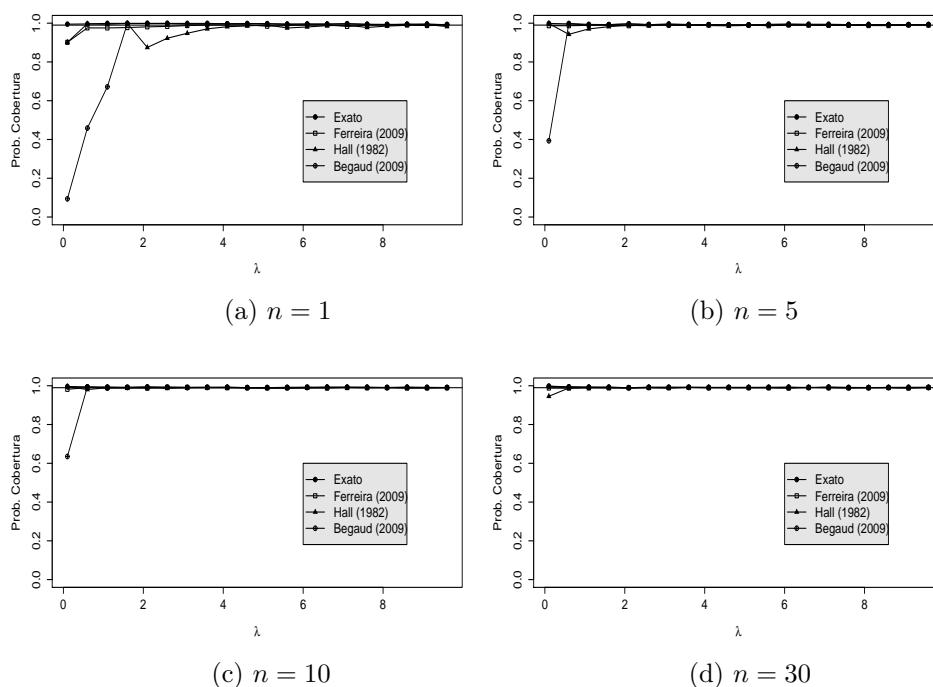
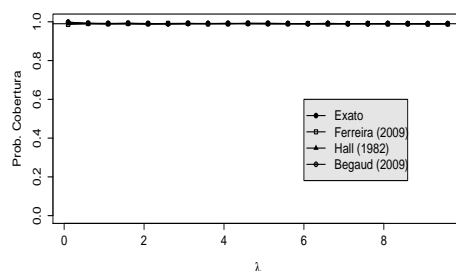


Figure 1 - Probabilidade de cobertura dos quatro intervalos estudados em $N = 10.000$ simulações Monte Carlo em função do valor de λ para diferentes valores de n : (a) $n = 1$, (b) $n = 5$, (c) $n = 10$ e (d) $n = 30$, considerando $\alpha = 0,01$.

Na Figura 2 está apresentada as coberturas dos 4 intervalos para $n = 50$, com $\alpha = 0,01$. Todos os intervalos apresentaram desempenho equivalentes para todos os valores de λ e foram considerados exatos, como pode ser observado pela não distinção entre linhas que identificam cada intervalo. Isso reforça o que foi

comentado anteriormente, que para grandes amostras, os intervalos assintóticos e exato tendem a se equivalerem, apresentando desempenhos ótimos em relação a probabilidade de cobertura.



(a) $n = 50$

Figure 2 - Probabilidade de cobertura dos quatro intervalos estudados em $N = 10.000$ simulações Monte Carlo em função do valor de λ para $n = 50$, considerando $\alpha = 0,01$.

As mesmas 4 situações das das Figuras 1 e 2, foram estudadas para $\alpha = 0,05$ (dados não apresentados). Comparando-se as duas situações, verificou-se o padrão de resposta foi praticamente o mesmo, acentuando-se alguns efeitos em alguns casos. Como exposto anteriormente, para grandes amostras, os intervalos assintóticos e exato tendem a se equivalerem, apresentando desempenho ótimo em relação a probabilidade de cobertura.

Na Tabela 3 estão apresentados os comprimentos médios dos quatro intervalos estudados, considerando diferentes valores de n e λ , com $\alpha = 0,05$. Essa situação equivale aos quatro casos da Tabela 2, em que foram avaliadas as probabilidades de cobertura. Verifica-se, para λ extremamente pequeno (0,01) e n pequeno ($n \leq 20$), que os comprimentos médios dos quatro intervalos tenderam a seguir uma determinada ordem, sendo o comprimento de intervalo de Ferreira o maior, seguido dos intervalos exato, Bégaud e Hall. O intervalo de Bégaud, nesse caso particular, apesar de ter tido o segundo menor comprimento de intervalo, apresentou probabilidade de cobertura próximo de 0%, para $n = 1$, que é o pior caso. Os demais intervalos apresentaram, nesse caso, cobertura próximas do valor nominal.

Para $n = 5$ e $\lambda = 0,01$, o mesmo padrão de resposta foi observado para o comprimento e probabilidade de cobertura dos quatro intervalos. Entretanto, o que ocorreu foi uma maior proximidade dos valores de cobertura, embora o intervalo de Hall (*IC3*), ainda se destacou com valores de comprimento bem menores do que os comprimentos dos demais intervalos. A cobertura do intervalo de Bégaud, continuou próxima de 0% com $n = 5$ (Tabela 2).

Com o aumento de n para 10 ou para 30 o mesmo padrão observado com $n = 1$ e 5, se verificou. Há de se ressaltar, no entanto, que os comprimentos dos intervalos

Table 3 - Comprimento médio dos quatro intervalos estudados (IC1 - Exato; IC2 - Ferreira (2009); IC3 - Hall (1982); IC4 - Bégau (2009)) em $N = 10.000$ para diferentes valores de n , considerando $\lambda = 0,01, 0,10, 0,50, 1,00, 5,00$ e $7,00$ com $\alpha = 0,05$

n	$\lambda = 0,01$				$\lambda = 0,10$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	3,7080	3,8588	0,9785	3,1581	3,8627	3,9955	1,2224	3,3875
5	0,7572	0,7855	0,2204	0,6528	0,9117	0,9223	0,4538	0,8587
10	0,3875	0,4007	0,1242	0,3378	0,5388	0,5346	0,3420	0,5228
15	0,2638	0,2719	0,0913	0,2329	0,4044	0,3966	0,2857	0,4002
20	0,2020	0,2076	0,0749	0,1815	0,3372	0,3279	0,2556	0,3372
25	0,1653	0,1700	0,0656	0,1500	0,2934	0,2837	0,2321	0,2950
30	0,1414	0,1444	0,0602	0,1282	0,2639	0,2537	0,2150	0,2659
50	0,0913	0,0924	0,0457	0,0859	0,1961	0,1874	0,1698	0,1985
100	0,0533	0,0530	0,0334	0,0521	0,1345	0,1284	0,1224	0,1349
1000	0,0134	0,0128	0,0122	0,0134	0,0402	0,0394	0,0392	0,0401
n	$\lambda = 0,50$				$\lambda = 1,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	4,5660	4,6179	2,2802	4,2898	5,3712	5,3315	3,3959	5,2393
5	1,4715	1,4222	1,1654	1,4718	1,9736	1,8853	1,7013	1,9766
10	0,9819	0,9380	0,8505	0,9908	1,3441	1,2838	1,2234	1,3486
15	0,7857	0,7497	0,7028	0,7896	1,0813	1,0362	1,0035	1,0825
20	0,6739	0,6437	0,6135	0,6738	0,9278	0,8918	0,8706	0,9287
25	0,5946	0,5688	0,5472	0,5974	0,8240	0,7940	0,7788	0,8253
30	0,5408	0,5183	0,5019	0,5418	0,7506	0,7249	0,7134	0,7506
50	0,4123	0,3973	0,3897	0,4134	0,5744	0,5579	0,5525	0,5752
100	0,2873	0,2791	0,2764	0,2873	0,4022	0,3934	0,3916	0,4018
1000	0,0886	0,0876	0,0876	0,0886	0,1250	0,1240	0,1239	0,1249
n	$\lambda = 5,00$				$\lambda = 7,00$			
	IC1	IC2	IC3	IC4	IC1	IC2	IC3	IC4
1	9,8409	9,4009	8,5273	9,9103	11,4560	10,9224	10,1914	11,5090
5	4,1267	3,9766	3,9009	4,1325	4,8486	4,6907	4,6269	4,8438
10	2,8764	2,7940	2,7673	2,8686	3,3830	3,2979	3,2754	3,3784
15	2,3299	2,2728	2,2583	2,3299	2,7460	2,6874	2,6752	2,7425
20	2,0110	1,9673	1,9578	2,0075	2,3707	2,3260	2,3180	2,3660
25	1,7942	1,7587	1,7519	1,7922	2,1143	2,0780	2,0723	2,1137
30	1,6347	1,6047	1,5996	1,6330	1,9270	1,8965	1,8922	1,9256
50	1,2597	1,2413	1,2389	1,2586	1,4873	1,4686	1,4666	1,4857
100	0,8862	0,8768	0,8759	0,8864	1,0473	1,0378	1,0371	1,0468
1000	0,2781	0,2771	0,2771	0,2781	0,3289	0,3279	0,3279	0,3288

de Ferreira e exato tenderam a se igualar e o comprimento do intervalo de Hall ainda continuou bem menor. As probabilidades de cobertura do intervalo de Bégau, só aproximaram do valor nominal de forma expressiva com $n = 30$ (Tabela 2), embora para $\lambda = 0,01$, isso só ocorreu a partir de n acima de 70 (Tabela 2). Assim, a superioridade do intervalo de Bégau em relação ao comprimento do intervalo, só pode ser considerada a partir de $n = 30$, para λ 's superiores a 0,10. Nesse caso, embora superior aos intervalos exato e de Ferreira, a diferença já não foi tão expressiva quanto a que ocorreu com $n = 1$.

Os quatro intervalos tenderam a apresentar o mesmo comprimento somente com $n = 1000$, indicando, a princípio, uma equivalência assintótica.

O padrão de resposta dos intervalos, quanto ao comprimento para $\lambda = 0,10$, foi o mesmo observado para $\lambda = 0,01$. O que se verificou foi uma aproximação dos comprimentos com n igual ou superior a 15. A partir desse ponto o intervalo de Ferreira passou a superar os intervalos exato e Bégau, embora as diferenças nos comprimentos dos intervalos não foram tão expressivas. Novamente, os comprimentos do intervalo de Hall foram bem menores do que os dos intervalos concorrentes. Essa diferença entre eles, só se dissipou com valores de n muito grandes, isto é, $n = 1000$.

Para $\lambda = 0,50, 1, 5$ e 7 (Tabela 3) os comprimentos de intervalos seguiram praticamente os mesmos padrões destacados para $\lambda = 0,01$ e $\lambda = 0,10$. As diferenças são mencionadas na sequência. Na medida que o valor de λ cresceu, os comprimentos dos intervalos tenderam a se igualar para um valor de n pequeno. Embora os valores dos comprimentos, em termos absolutos, também aumentem com o aumento de λ , os valores relativos são cada vez menores. O intervalo de Ferreira (*IC2*) apresentou menor comprimento do que o exato para todos os valores de n a partir de $\lambda \geq 1$, embora para grandes valores de n os comprimentos foram praticamente iguais. Os comprimentos do *IC3* de Hall, foram sempre menores do que os demais intervalos. No entanto, a superioridade deixa de ser tão evidente à partir de $n = 5$, para valores de λ grandes. O *IC4* de Bégau, quase sempre apresentou comprimentos próximos aos obtidos para o *IC1*, intervalo exato.

Os desempenhos das três aproximações estudadas nesse trabalho foram bastantes variados e, em alguns casos, inesperados. As probabilidades de coberturas foram muitas vezes inferiores aos valores nominais do coeficiente de confiança e sem um padrão claro para explicar os resultados insatisfatórios. Por exemplo, o intervalo de Hall apresentou, para $\lambda = 0,01$, com n variando de 2 a 30, cobertura abaixo do nominal de 99%.

Nesse mesmo caso de $\lambda = 0,01$, o intervalo de Ferreira apresentou ciclos de queda e recuperação da probabilidade de cobertura com a variação de n , o que também não se conseguiu explicar. Embora, com comportamento mais regular, o intervalo de Bégau apresentou baixo desempenho em combinações com pequenos valores de n e λ .

Dessa forma, os desempenhos superiores dessas aproximações, em algumas situações, em relação aos comprimentos dos intervalos, que são menores, devem ser vistos com reserva. Isso porque a probabilidade de cobertura dos mesmos, em geral, foram inferiores aos valores nominais, para pequenos valores de λ e n . Em geral, se propõe um intervalo aproximado, quando os comprimentos dos intervalos exatos são grandes. Como os desempenhos dos intervalos aproximados e exato só se equivalem em combinações de grandes valores de n e λ , então não há vantagem para se utilizar tais aproximações. Portanto, como conclusão final, recomenda-se preferencialmente o uso do intervalo exato em todos os casos. As aproximações, no entanto, podem ser utilizadas em algumas circunstâncias teóricas, que se necessite derivar algum outro resultado, em que seja impossível utilizar resultados exatos.

Conclusões

As principais conclusões desse trabalho são:

Uma solução explícita para o intervalo de Hall (1982) foi apresentada no presente trabalho concomitantemente com a avaliação de seu desempenho e, assim, disponibilizar para a comunidade científica uma expressão simples de ser utilizada.

O intervalo de confiança aproximado de Bégau et al. (2005) foi generalizado para amostras com $n \geq 1$ e avaliado por simulação Monte Carlo.

Os três intervalos assintóticos apresentam desempenho equivalente em relação a probabilidade de cobertura ao intervalo exato a partir de $\lambda \geq 5$ e $n > 1$. No caso de $n = 1$ e $\lambda \geq 5$, o intervalo de Bégau et al. (2005) foi o que desempenhou melhor.

Em relação ao comprimento dos intervalos, para $\lambda \geq 5$ e $n > 1$, os desempenhos dos intervalos de confiança seguiram a seguinte ordem; do melhor (menor comprimento) para o pior (maior comprimento): Intervalo de Hall (1982), Ferreira (2009), Bégau et al. (2005) e exato.

Como em pequenas amostras para pequenos valores de λ as aproximações tiveram baixo desempenho e o valor de λ é desconhecido, recomenda-se a utilização do intervalo exato, exceto na derivação de algum resultado em que a teoria exata seja impossível de ser aplicada.

Agradecimentos

Os autores agradecem ao *CNPq* pelo suporte financeiro para o trabalho.

CORRÊA, C. B. V.; FERREIRA, D. F. Performance of confidence intervals for the mean of Poisson populations evaluated by Monte Carlo simulations. *Rev. Bras. Biom.*, São Paulo, v.31, n.3, p.423-440, 2013.

■ **ABSTRACT:** *Many approximations to confidence intervals for the parameter λ of a Poisson population exist in the literature. The present work was made with the main objective to evaluate, by simulation, the proprieties of asymptotes procedures to confidence interval for the parameter λ of Poisson. Three asymptotic and the exact intervals were studied. The simulations were performed in R statistical software. The specific objectives were reach an explicit solution for Hall (1982)'s interval and generalize the Bégau et al. (2005)'s interval for the case where $n > 1$. As proposed, an explicit solution for the interval of Hall (1982) has been achieved. The approximated confidence interval of Bégau et al. (2005) was generalized to sample sizes greater than 1. The three asymptotic intervals have shown equivalent performance to the exact interval with respect to the coverage probability to $\lambda \geq 5$ and $n > 1$. With $n = 1$ and $\lambda \geq 5$, the Bégau et al. (2005) interval show the best performance. Regards to the interval lengths, for $\lambda \geq 5$ and $n > 1$, the performance of the confidence intervals followed the order, from the best (shorter) to the worst (longest) case: Hall (1982), Ferreira (2009), Bégau et al. (2005) and exact. As in small samples, for small values of λ , the approximations had low performance and λ is unknown, it is recommended to use the exact interval, except when some results should be built and exact theories are not available and are difficult or impossible to construct.*

- **KEYWORDS:** Normal approximation of Poisson; interval length; coverage probability, Monte Carlo.

References

- AHRENS, J. H.; DIETER, U. Computer generation of poisson deviates from modified normal distributions. *ACM Transactions on Mathematical Software*, v. 8, p. 163–179, 1982.
- BÉGAUD, B. et al. An easy to use method to approximate poisson confidence limits. *European Journal of Epidemiology*, v. 20, p. 213–216, 2005.
- BLYTH, C. R.; HUTCHINSON, D. W. Table of neyman-shortest unbiased confidence intervals for the poisson parameter. *Biometrika*, v. 48, p. 191-134, 1961.
- FERREIRA, D. F. *Estatística Básica*. 1^a. ed. Lavras: Editora UFLA, 2009. 664 p.
- GARWOOD, F. Fiducial limits for the poisson distribution. *Biometrika*, v. 28, n. 3-4, p. 437–442, 1936.
- HALL, P. Improving the normal approximation when constructing one-sided confidence intervals for binomial or poisson parameters. *Biometrika*, v. 69, n. 3, p. 647-652, 1982.
- PATIL, V.; KULKARNI, H. Comparison of confidence intervals for the poisson mean: Some new aspects. *REVSTAT – Statistical Journal*, v. 10, n. 2, p. 211–227, 2012.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2013. Disponível em: <<http://www.R-project.org/>>.

Received in 28.08.2013.

Approved after revised in 06.11.2013.

APÊNDICE

Funções R

```
# Função para obter o intervalo de confiança exato de uma
# amostra aleatória independente e identicamente distribuida,
# (X_1, X_2, ..., X_n), da Poisson(lamb)
IcPoiExact <- function(x, alpha = 0.05)
{
  k <- sum(x)
  n <- length(x)
  if (k == 0) ll <- 0 else
  ll <- qchisq(alpha / 2, 2 * k) / (2 * n)
  ul <- qchisq(1 - alpha / 2, 2 * k + 2) / (2 * n)
  return(list(ci.lower=ll,ci.upper=ul,confidence.prob=1-alpha))
}

# Função para obter o intervalo de confiança Ferreira (2009)
# de uma amostra aleatória independente e identicamente
# distribuida, (X_1, X_2, ..., X_n), da Poisson(lamb)
IcPoiApr1 <- function(x, alpha = 0.05)
{
  z <- qnorm(1 - alpha/2)
  n <- length(x)
  Xbar <- mean(x)
  ll <- -(z^2/n+2*Xbar-z/n^0.5*(z^2 / n + 4 * Xbar)^0.5) / 2
  ul <- -(z^2/n+2*Xbar+z/n^0.5*(z^2 / n + 4 * Xbar)^0.5) / 2
  return(list(ci.lower=ll,ci.upper=ul,confidence.prob=1-alpha))
}

# Função para obter o intervalo de confiança Hall (1982) de uma
# amostra aleatória independente e identicamente distribuida,
# (X_1, X_2, ..., X_n), da Poisson(lamb)
IcPoiApr2 <- function(x, alpha = 0.05)
{
  z <- qnorm(1 - alpha/2)
  n <- length(x)
  Xbar <- mean(x)
  if (sum(x) == 0) ll <- 0 else
  ll <- Xbar - z * (Xbar / n)^0.5 - (1 - z^2) / (3 * n)
  ul <- Xbar + z * (Xbar / n)^0.5 - (1 - z^2) / (3 * n)
  return(list(ci.lower=ll,ci.upper=ul,confidence.prob=1-alpha))
}
```

```

# Função para obter o intervalo de confiança de Begaud et al. (2005)
# modificado para n >= 1 de uma amostra aleatória independente
# e identicamente distribuída, (X_1, X_2, ..., X_n) da
# Poisson(lamb)
IcPoibigaude <- function(x, alpha = 0.05)
{
  z <- qnorm(1 - alpha/2)
  n <- length(x)
  Xbar <- mean(x)
  ll <- n * (((Xbar + 0.02/n)/n)^0.5 - z/(2*n))^2
  ul <- n * (((Xbar + 0.96/n)/n)^0.5 + z/(2*n))^2
  return(list(ci.lower=ll,ci.upper=ul,confidence.prob=1-alpha))
}

```